Apéndice B

Bases de datos con variable consecuente difuso

En este apéndice se describen las bases de datos con consecuente difuso utilizadas en la parte experimiental de esta memoria. Debemos destacar que todas las bases de datos se han generado por el autor de este trabajo, excepto la base RICE que se ha obtenido del trabajo [119]. Sobre todas, se han construido cinco particiones de entrenamiento y prueba para evitar el sesgo de los resultados obtenidos. Todas las bases de datos, se han discretizado usando siete etiquetas difusas uniformemente repartidas sobre el referencial de cada variables, excepto en el caso de RICE que se han considerado únicamente cinco.

A continuación pasamos a describir cada una de las bases de datos.

- **Base de datos SIN**

  Descripción: Esta base de datos contiene 1000 ejemplos obtenidos por la aplicación de la función $\sin(x)$ donde $x \in [0, 1]$.

- **Base de datos PARABOLA Descripción:** Esta base de datos contiene 1000 ejemplos obtenidos por la aplicación de la función x^2 donde $x \in [0, 1]$.

- **Base de datos 3DPARABOLA**

  Descripción: Esta base de datos contiene 1000 ejemplos obtenidos por la aplicación de la función $f(x, y) = x^2 + y^2$ donde $x, y \in [-5, 5]$.

- **Base de datos MONTE**

  Descripción: Esta base de datos contiene 1000 ejemplos obtenidos por la aplicación de la función $f(x, y) = (cos^2(x) + cos^2(y))^4$ y que se muestra la figura B.4. La base de datos contiene dos variables predictivas $x, y \in [-1, 1]$.

- **Base de datos PENDULO**

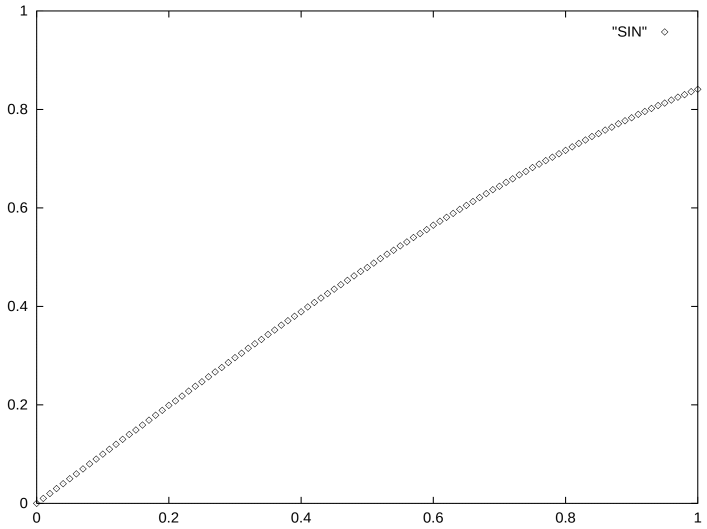

Figura B.1: Representación de la base de datos SIN

Descripción: El problema del péndulo es un problema tradicional de control y consiste en que, dado un conjunto de ejemplos que describe el comportamiento del sistema, deseamos aprender el conjunto de reglas que permiten controlarlo. Si asumimos que el valor de $|\theta| << 1$ (radianes) la equación diferencial no lineal que describe el comportamiento del péndulo es

$$m\frac{l^2}{3}\frac{d^2\theta}{dt^2} = \frac{l}{2}(-f + m\ g\ sin\theta - k\frac{d\theta}{dt})$$

donde $k\frac{d\theta}{dt}$ es una aproximación de la fuerza de fricción.

Por tanto, el sistema se puede describir a través de dos variables de estado θ (ángulo) y ω (velocidad angular) y la variable de control f (fuerza). Para la simulación real se ha considerado un péndulo de 5 kg de peso y 5 m de largo, al que se han aplicado fuerzas sobre el centro de gravedad cada 10 ms. Con estos parámetros, el universo de discurso de las variables son los siguientes:

$$\theta \in [-0.277, 0.277] \quad \omega \in [-0.458, 0.458] \quad f \in [-1593, 1593]$$

De forma experimental, hemos obtenido dos conjuntos de ejemplos bajo las anteriores condiciones, el primero de ellos tiene 213 ejemplos y se ha usado para aprender (conjunto de entrenamiento) y el segundo contiene 125 ejemplos y se ha utilizado para comprobar el comportamiento de las reglas aprendidas (conjunto de test). Estos conjuntos de ejemplos se han obtenido de las ecuaciones anteriores teniendo en cuenta dos condiciones iniciales diferentes:

a) θ=-0.277 y ω=0.

b) θ=0.277 y ω=0.

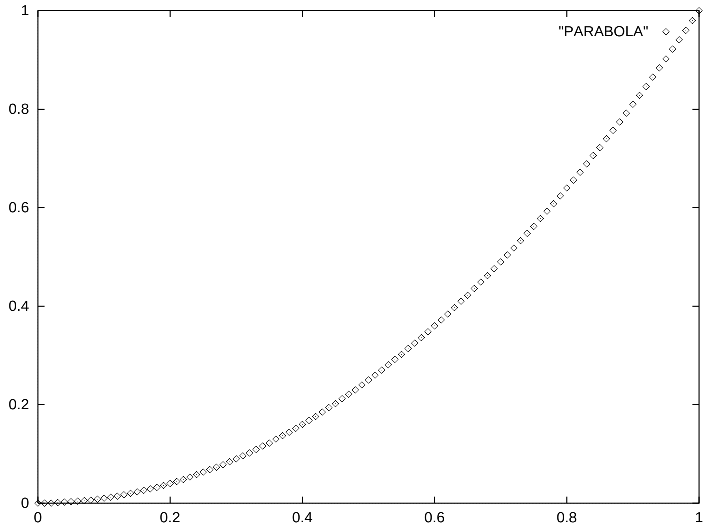

Figura B.2: Representación de la base de datos PARABOLA

El primer paso para aprender consiste en discretizar el rango de las variables mediante etiquetas difusas. La figura B.5 muestras las etiquetas construidas para cada variable, mientras que la figura B.6 muestra la función que describe el comportamiento del péndulo.

- **Base de datos RICE**

  Descripción: Esta base de datos se ha obtenido a través del trabajo [119]. Los datos se han extraído de un problema real. La base de datos RICE tiene cinco variables de entrada y una salida cuyos valores estan asociados con las siguientes medidas subjetivas: sabor, apariencia, gusto, viscosidad y dureza. La variable de salida establece una evaluación general del arroz. La base de datos contiene 105 ejemplos que muestran 105 tipos diferentes de arroz.

  Para este problema, se han considerado cinco particiones en cada variable para construir el dominio difuso asociado tanto a las variables de entrada, como a la variable de salida.

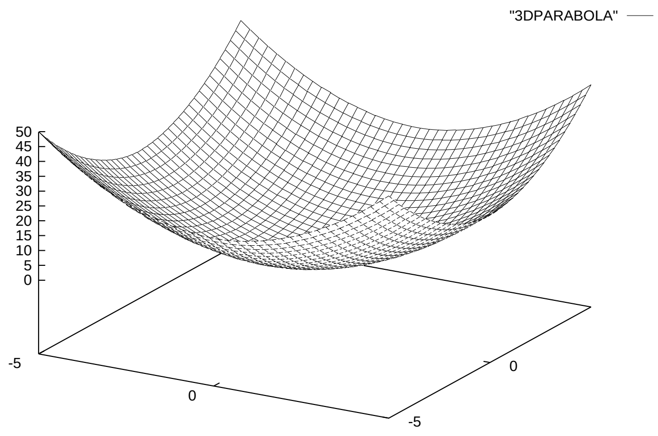

Figura B.3: Representación de la base de datos 3DPARABOLA

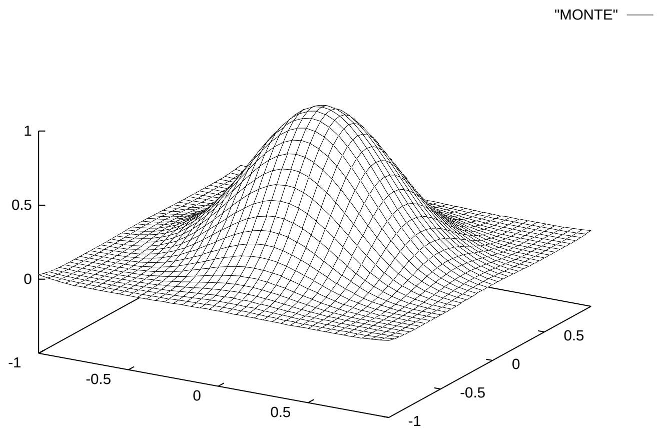

Figura B.4: Representación de la base de datos MONTE

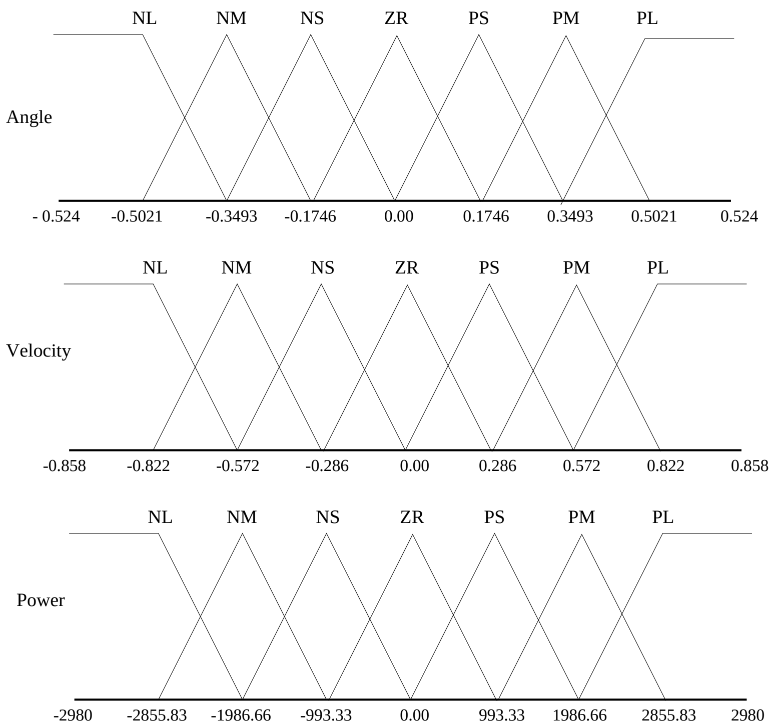

Figura B.5: El problema del péndulo

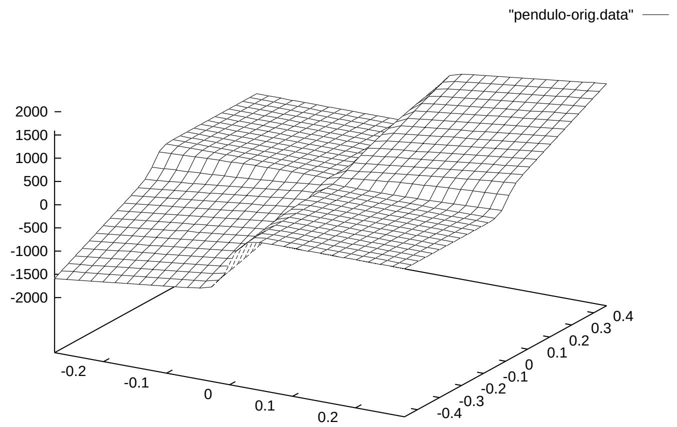

Figura B.6: Superficie original del problema del péndulo